

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-069290

(43)Date of publication of application : 09.03.1999

(51)Int.Cl.

H04N 5/91
H04N 5/262
H04N 5/76
H04N 5/7826

(21)Application number : 09-220260

(71)Applicant : NIPPON TELEGR & TELEPH CORP
<NTT>

(22)Date of filing : 15.08.1997

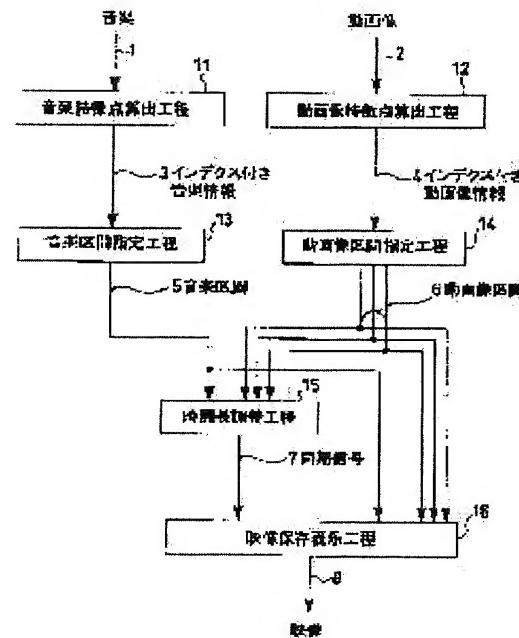
(72)Inventor : NIIKURA YASUMASA
MINAMI KENICHI
AKUTSU AKITO
TONOMURA YOSHINOBU

(54) VIDEO PRODUCTION METHOD AND DEVICE AND RECORD MEDIUM RECORDED THIS METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To produce a good video image, in the case that a source is edited and the video image is produced and at the time of synchronization processing of music and dynamic image and synchronization.

SOLUTION: First, each characteristic point is calculated from a music 1 and a dynamic image which is a source at a music characteristic point calculation process 11 and a dynamic image characteristic calculation process 12. Then a music block 5 and a dynamic image block 6 which the user desires as an object are designated by a music designation process 13 and a dynamic image designation process 14. Then the dynamic image block 6 is combined with the music block 5 to adjust a time length required to configure a video image at a time length adjustment process 15. The video image 8 with a music being a combination of the music and the dynamic image whose time length is adjusted is stored and displayed by a video storage display process 16. Thus, a user selects the music 1 and the dynamic image 2 of the designated source, and each characteristic point is used to make a combination such that an effective video image is obtained through time synchronization, thereby producing an effective video image.



LEGAL STATUS

[Date of request for examination] 22.12.2000

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平11-69290

(43)公開日 平成11年(1999)3月9日

(51) Int.Cl.⁶
H 0 4 N 5/91
5/262
5/76
5/7826

識別記号

F I
H 0 4 N 5/91
5/262
5/76
5/7826

N
Z
A

審査請求 未請求 請求項の数25 O.L (全 16 頁)

(21)出願番号 特願平9-220260

(22)出願日 平成9年(1997)8月15日

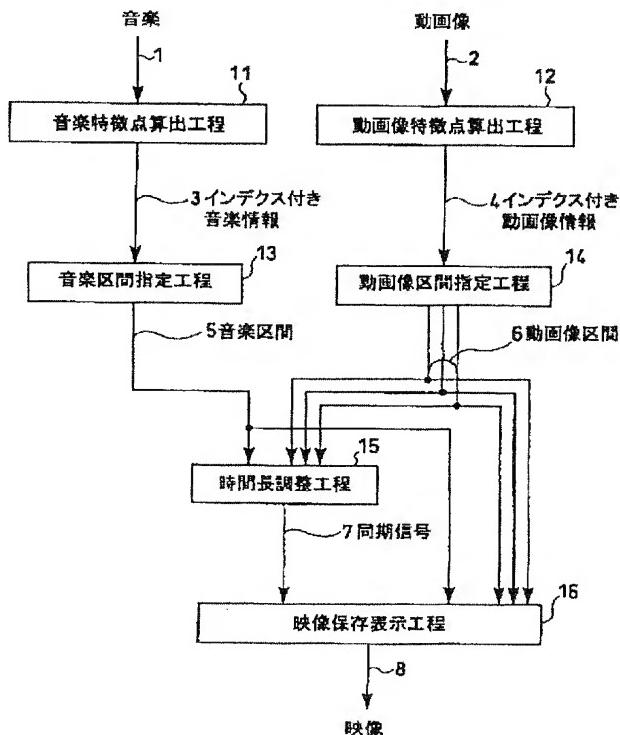
(71)出願人 000004226
日本電信電話株式会社
東京都新宿区西新宿三丁目19番2号
(72)発明者 新倉 康巨
東京都新宿区西新宿3丁目19番2号 日本
電信電話株式会社内
(72)発明者 南 憲一
東京都新宿区西新宿3丁目19番2号 日本
電信電話株式会社内
(72)発明者 阿久津 明人
東京都新宿区西新宿3丁目19番2号 日本
電信電話株式会社内
(74)代理人 弁理士 志賀 富士弥 (外1名)
最終頁に続く

(54)【発明の名称】 映像制作方法及び装置及びこの方法を記録した記録媒体

(57)【要約】

【課題】 素材を編集し映像を制作する場合、音楽と動画像の同期処理及び同期の際に良い映像ができる映像制作方法および装置を実現する。

【解決手段】 まず、素材の音楽1と動画像2から各々の特徴点を音楽特徴点算出工程11と動画像特徴点算出工程12により算出する。次に、利用者が対象としたい音楽区間5と動画像区間6を、各々音楽指定工程13と動画像指定工程14で指定する。そして、音楽区間5と動画像区間6を組み合わせて映像を構成するために必要な時間長を時間長調整工程15である。この時間長を調節した音楽と動画像を組み合せた音楽付き映像8を、映像保存表示工程16により保存、表示する。以上により、利用者が指示した素材の音楽1と動画像2とを選択し、各々の特徴点を用いて時間同期をとって効果的な映像となるように組み合わせを行うことで、効果的な映像の制作を可能とする。



【特許請求の範囲】

【請求項1】 動画像と音楽を組み合わせて、音楽付き映像を制作する映像制作方法において、利用者が対象としている音楽区間を指定する音楽指定工程と、利用者が対象としている動画像区間を指定する動画像指定工程と、指定された該音楽区間に、指定された該動画像区間を組み合わせて映像を構成するために必要な時間長をあわせる時間長調整工程と、時間長を調整した該音楽区間と該動画像区間を組み合わせて構成された映像を保存、表示する映像保存表示工程と、を具備することを特徴とする映像制作方法。

【請求項2】 前記動画像指定工程の前に、動画像区間を指定する際に、動画像からあらかじめ定められた一種類乃至は複数種類の特徴的なイベントが発生したことを示す特徴点を算出しておく動画像特徴点算出工程を具備する、

ことを特徴とする請求項1に記載の映像制作方法。

【請求項3】 前記音楽指定工程の前に、音楽区間を指定する際に、音楽からあらかじめ定められた一種類乃至は複数種類の特徴的なイベントが発生したことを示す特徴点を算出しておく音楽特徴点算出工程を具備する、ことを特徴とする請求項1または2に記載の映像制作方法。

【請求項4】 前記動画像指定工程では、動画像区間の指定に際して、あらかじめ抽出された該特徴点を基準に、一つ乃至複数の動画像区間を選択する、

ことを特徴とする請求項2または3に記載の映像制作方法。

【請求項5】 前記時間長調整工程では、該音楽区間に含まれる特徴点を、時間同期させる動画像の切り替え基準点として設定し、該切り替え基準点毎に動画像を切り替える、

ことを特徴とする請求項3または4に記載の映像制作方法。

【請求項6】 前記時間長調整工程では、該動画像区間に含まれる特徴点を、時間同期させる音楽の切り替え基準点として設定し、該切り替え基準点毎に音楽を切り替える、

ことを特徴とする請求項2、3、4のいずれか1項に記載の映像制作方法。

【請求項7】 前記時間長調整工程では、該音楽区間に含まれる特徴点を基に設定された該切り替え基準点毎に該動画像区間を切り替える場合に、該動画像区間の任意の一部を削除することによって該動画像区間の時間長を調整し、該音楽区間との同期を行う、

ことを特徴とする請求項5に記載の映像制作方法。

【請求項8】 前記時間長調整工程では、該動画像区間

に含まれる特徴点を基に設定された該切り替え基準点毎に該音楽区間を切り替える場合に、該音楽区間の任意の一部を削除することによって該音楽区間の時間長を調整し、該動画像区間との同期を行う、

ことを特徴とする請求項6に記載の映像制作方法。

【請求項9】 前記時間長調整工程では、該動画像区間の時間短縮を目的とした削除を行う場合で、該動画像区間の削除を禁止する区間を設ける、

ことを特徴とする請求項7に記載の映像制作方法。

10 【請求項10】 前記時間長調整工程では、該音楽区間の時間短縮を目的とした削除を行う場合において、該音楽区間の削除を禁止する区間を設ける、

ことを特徴とする請求項8に記載の映像制作方法。

【請求項11】 前記時間長調整工程では、該動画像区間における時間短縮を目的とした削除を禁止する区間を設ける場合に、その区間を動画像区間に含まれる特徴点を基に設定する、

ことを特徴とする請求項9に記載の映像制作方法。

20 【請求項12】 前記時間長調整工程では、該音楽区間における時間短縮を目的とした削除を禁止する区間を設ける場合に、その区間を音楽区間に含まれる特徴点を基に設定する、

ことを特徴とする請求項10に記載の映像制作方法。

【請求項13】 請求項1乃至12のいずれか1項に記載の映像制作方法における各工程の手順をコンピュータに実行させるための処理プログラムを、該コンピュータが読み取り可能な媒体に記録した、

ことを特徴とする映像制作方法を記録した記録媒体。

30 【請求項14】 動画像と、音楽を組み合わせて、音楽付き映像を制作する映像制作装置において、利用者が対象としている音楽区間を指定する音楽指定部と、

利用者が対象としている動画像区間を指定する動画像指定部と、

指定された該音楽区間に、指定された該動画像区間を組み合わせて映像を構成するために必要な時間長をあわせる時間長調整部と、

時間長を調節した該音楽区間と該動画像区間を組み合わせて構成された映像を保存、表示する映像保存表示部と、

を具備することを特徴とする映像制作装置。

【請求項15】 前記動画像区間を指定する際に、動画像からあらかじめ定められた一種類乃至は複数種類の特徴的なイベントが発生したことを示す特徴点を算出しておく動画像特徴点算出部を新たに具備する、

ことを特徴とする請求項14に記載の映像制作装置。

【請求項16】 前記音楽区間を指定する際に、音楽からあらかじめ定められた一種類乃至は複数種類の特徴的なイベントが発生したことを示す特徴点を算出しておく音楽特徴点算出部を新たに具備する、

ことを特徴とする請求項14または15に記載の映像制作装置。

【請求項17】 前記動画像指定部は、動画像区間の指定に際して、あらかじめ抽出された該特徴点を基準に、一つ乃至複数の動画像区間を選択するものである、ことを特徴とする請求項15または16に記載の映像制作装置。

【請求項18】 前記時間長調整部は、該音楽区間に含まれる特徴点を、時間同期させる動画像の切り替え基準点として設定し、該切り替え基準点毎に動画像を切り替えるものである、

ことを特徴とする請求項16または17に記載の映像制作装置。

【請求項19】 前記時間長調整部は、該動画像区間に含まれる特徴点を、時間同期させる音楽の切り替え基準点として設定し、該切り替え基準点毎に音楽を切り替えるものである、

ことを特徴とする請求項15, 16, 17のいずれか1項に記載の映像制作装置。

【請求項20】 前記時間長調整部は、該音楽区間に含まれる特徴点を基に設定された該切り替え基準点毎に該動画像区間を切り替える場合に、該動画像区間の任意の一部を削除することによって該動画像区間の時間長を調整し、該音楽区間との同期を行うものである、

ことを特徴とする請求項18に記載の映像制作装置。

【請求項21】 前記時間長調整部は、該動画像区間に含まれる特徴点を基に設定された該切り替え基準点毎に該音楽区間を切り替える場合に、該音楽区間の任意の一部を削除することによって該音楽区間の時間長を調整し、該動画像区間との同期を行うものである、

ことを特徴とする請求項19に記載の映像制作装置。

【請求項22】 前記時間長調整部は、該動画像区間の時間短縮を目的とした削除を行う場合で、該動画像区間の削除を禁止する区間を設ける機能を具備する、

ことを特徴とする請求項20に記載の映像制作装置。

【請求項23】 前記時間長調整部は、該音楽区間の時間短縮を目的とした削除を行う場合において、該音楽区間の削除を禁止する区間を設ける機能を具備する、

ことを特徴とする請求項21に記載の映像制作装置。

【請求項24】 前記時間長調整部は、該動画像区間ににおける時間短縮を目的とした削除を禁止する区間を設ける場合に、その区間を動画像区間に含まれる特徴点を基に設定するものである、

ことを特徴とする請求項22に記載の映像制作装置。

【請求項25】 前記時間長調整部は、該音楽区間ににおける時間短縮を目的とした削除を禁止する区間を設ける場合に、その区間を音楽区間に含まれる特徴点を基に設定するものである、

ことを特徴とする請求項23に記載の映像制作装置。

【0001】

【発明の属する技術分野】本発明は、動画像と音楽を組み合わせて映像を制作する方法および装置に関するものである。その際に問題となる動画像と音楽の時間同期を自動的に行うため、音楽及び動画像から抽出した特徴点から同期に用いる基準点を設定し、その基準点に従って一部を削除する等の操作を行って同期を行う。ただし、音楽または動画像にある特徴的な内容が含まれている場合には、一部を削除するなどの操作を行わない区間を動画像や音楽に与えることで、重要なコンテンツを残したまま、時間同期が自動的に行われた映像を制作することができ、映像の自動制作といった分野への応用が考えられる。

【0002】

【従来の技術】ディジタル技術の進歩により、映像を放送局やプロダクション以外の一般の人々が気軽に扱えるようになった。一般の人々は各自の興味のある対象を小型ビデオカメラで撮影したり、ネットワークでそれに類する映像を入手したりして、映像作品を制作することができる時代になっている。

【0003】一方で、こうした映像を視聴する側には、膨大な映像の中から自分に有益な映像を選択する必要が生じる。そこで、視聴者の目に付きやすいように内容を要約し、かつ宣伝するための映像を提供する必要がある。また、本編にあたる映像作品においても、その冒頭だけを見て、視聴者が飽きて視聴を止めてしまうことを防ぐために、興味を惹くような映像を冒頭や途中、乃至は全編に渡って必要とする。

【0004】こうした映像は、音楽が付加してあった

30 たり、映像作品の一部を名場面としてつないだりなどの工夫が凝らされているのが一般的である。これらの映像を制作するためには、良い動画像素材と良い音楽素材を、最適に組み合せる必要がある。

【0005】こうした映像は、現在のテレビ番組や映画における次回予告、上映予告、番組宣伝、または、オープニング映像と同等の利用価値が見出されることもある。これらは、本編よりも十分に短く、かつ、本編の魅力を十分に伝える効果が求められる。これらの映像の作成方法は、様々なノウハウの積み重ねで作られているが、なかでも、代表的な制作方法として、映像本編の内容の雰囲気や魅力を効果的にあらわすテーマを表す音楽と、本編で用いられた、または、深く関連する映像の一部を組み合わせて制作する方法である。こうすることで本編の一部がわかり、かつ、テーマ音楽の利用によって映像本編の雰囲気を伝える効果を有する。なお、便宜上、テーマ音楽に動画像情報を付加することで制作された映像をテーマ音楽映像と呼ぶ。

【0006】上述のテーマ音楽映像を初めとする動画像と音楽との組み合せによる映像の作成には、まず素材として利用する音楽と動画像が重要となる。次に、音楽と

動画像の時間長が一致していることが最低条件である。さらに、音楽のメロディの変化や音量の変化と共にシーンが変化する等の工夫が行われている方がより完成度の高い映像になることが知られている。

【0007】例えば、テーマ音楽映像の制作において、従来の映像制作プロダクション等では、全てを人手を用いて制作、編集している。しかし、これは本編に加えて新たに一本の作品を構成することと等しい。

【0008】そこで、テーマ音楽映像の制作を目的の一つとして、映像中の音楽、人の声、等を分類し、映像中のより重要な部分を自動的に抽出して映像のダイジェストを制作する試みが(Enhanced Video Handling based on Audio Analysis: K. Minami, A. Akutsu, H. Hamada, Y. Tonomura: Proceedings of the IEEE Multimedia Computing and Systems 1997, p. 219)によって行われている。しかし、こうした自動処理によって作成された映像は、最適な動画像や音楽が素材として選択されているとは言い難く、次回予告や番組宣伝などの目的を十分満たしているとは言い難い。すなわち、素材の選択には、内容把握の能力が必要となるため、自動化が困難な面がある。

【0009】テーマ音楽映像の場合には、その素材となる動画像は、本編の動画像から選択すれば良いだけなので、従来の映像作品の制作とは異なって素材を集めるという労力は軽減され人手を介しても十分容易に行うことができる。しかし、それら素材を編集する作業は従来と同様の困難が伴う。編集作業の労力を軽減する目的で作成されたノンリニア編集機は、頭出しや反復して表示するなどの編集作業時の労力を軽減する目的には対応しているが、音楽と動画像に工夫を加えながら自動的に同期する機能はない。

【0010】

【発明が解決しようとする課題】上述したように動画像と音楽を組み合せによる映像の作成のための従来技術の問題点は、映像の作成に用いる良い素材を自動抽出できないこと、と、素材としての音楽及び動画像を簡単に同期させる機能が不十分であることの2点があげられる。しかし、素材の選択方法は非常に困難を伴うため、自動化することは困難である。従って、本発明ではこれは取り上げない。

【0011】本発明は、前述した問題点のうちの素材を編集し映像を制作する処理における問題点である音楽と動画像の同期処理及び同期の際に良い映像ができるような工夫を含めた編集処理、すなわち自動同期を行う映像制作方法および装置を実現することを課題とし、その課題を解決するためのものである。

【0012】

【課題を解決するための手段】前述の課題解決のため、

本発明は、以下の(1)～(13)の発明によって利用者が音楽と動画像をあらかじめ選択した後、音楽と動画像からより効果的な映像を作成しつつ、自動的に音楽と動画像の同期を行なうことを可能にする。

【0013】(1)の発明は、動画像と音楽を組み合わせて、音楽付き映像を制作する映像制作方法または装置において、利用者が対象とした音楽区間を指定する音楽指定工程または音楽指定部と、利用者が対象とした動画像区間を指定する動画像指定工程または動画像指定部と、指定された該音楽区間に、指定された該動画像区間を組み合わせて映像を構成するために必要な時間長をあわせる時間長調整工程または時間長調整部と、時間長を調節した該音楽区間と該動画像区間を組み合わせて構成された映像を保存、表示する映像保存表示工程または映像保存表示部と、を具備することを特徴とする。この発明により、音楽付き映像を制作する際の素材となる音楽と動画像とを選択し、より効果的な映像となるように時間同期をとてこれらを組み合わせて構成することで、効果的な映像の制作、表示、保存を可能とする。

【0014】(2)の発明は、(1)の発明の映像制作方法または装置において、動画像区間の指定するため、動画像からあらかじめ定められた一種類乃至は複数種類の特徴的なイベントが発生したことを示す特徴点を算出しておく動画像特徴点算出工程または動画像特徴点算出部を具備することを特徴とする。なお、動画像特徴点算出工程は動画像指定工程の前に具備する。

【0015】(3)の発明は、(1)または(2)の発明の映像制作方法または装置において、音楽区間の指定する際に、音楽からあらかじめ定められた一種類乃至は複数種類の特徴的なイベントが発生したことを示す特徴点を算出しておく音楽特徴点算出工程または音楽特徴点算出部を具備することを特徴とする。なお、音楽特徴点算出工程は、音楽指定工程の前に具備する。

【0016】これらの(2)、(3)の発明により音楽と動画像からなる映像を制作する際に、音楽及び動画像の両素材をより効果的な同期をとる際に必要な基準点に用いる特徴的なイベントが発生したことを示す特徴点を算出できる。また、これらの特徴点は素材に用いる情報区間の選択にも有効利用できるため、効果的な映像制作を可能とする。

【0017】(4)の発明は、(2)または(3)の発明の映像制作方法または装置における動画像指定工程または動画像指定部において、動画像区間の指定に際して、あらかじめ抽出された該特徴点を基準に、一つ乃至複数の動画像区間を選択することを特徴とする。この発明により、複数の動画像を抽出することで、映像作品の様々な動画像の一部ずつを抽出して利用することができる。

【0018】(5)の発明は、(2)～(4)のいずれかの発明の映像制作方法または装置における時間長調整工程または時間長調整部において、該音楽区間に含まれ

る特徴点を、時間同期させる動画像の切り替え基準点として設定し、該切り替え基準点毎に動画像を切り替えること、または、該動画像区間に含まれる特徴点を、時間同期させる音楽の切り替え基準点として設定し、該切り替え基準点毎に音楽を切り替えることを特徴とする。この発明により、音楽から抽出された特徴点を動画像切り替えの為の基準点に用いるため、音楽の変化にあわせて動画像のシーンチェンジを発生させることを可能にし、完成度の高い映像の制作を可能にしている。

【0019】(6)の発明は、(5)の発明の映像制作方法または装置における時間長調整工程または時間長調整部において、該音楽区間に含まれる特徴点を基に設定された該切り替え基準点毎に該動画像区間を切り替える場合に、該動画像区間の任意の一部を削除することによって該動画像区間の時間長を調整し、該音楽区間との同期を行うこと、または、該動画像区間に含まれる特徴点を基に設定された該切り替え基準点毎に該音楽区間を切り替える場合に、該音楽区間の任意の一部を削除することによって該音楽区間の時間長を調整し、該動画像区間との同期を行うこと、を特徴とする。この発明により、映像作成時の時間同期を行う際に、音楽からのインデックスにあわせて、動画像の一部を削除し、調整することを可能にしている。

【0020】(7)の発明は、(6)の発明の映像制作方法または装置における時間長調整工程または時間長調整部において、該動画像区間の時間短縮を目的とした削除を行う場合、該動画像区間の削除を禁止する区間を設ける機能を有すること、または、該音楽区間の時間短縮を目的とした削除を行う場合、該音楽区間の削除を禁止する区間を設ける機能を有すること、を特徴とする。

【0021】(8)の発明は、(7)の発明の映像制作方法または装置における時間長調整工程または時間長調整部において、該動画像区間における時間短縮を目的とした削除を禁止する区間を設ける場合に、その区間を動画像区間に含まれる特徴点を基に設定すること、または、該音楽区間における時間短縮を目的とした削除を禁止する区間を設ける場合に、その区間を音楽区間に含まれる特徴点を基に設定すること、を特徴とする。

【0022】これらの(7)、(8)の発明により、時間同期のための時間短縮を目的とした動画像の一部の削除において、より重要と思われるカメラワークや動き被写体が存在している動画像の削除を防ぐことが可能となる。従って、より効果的な映像を作成できる。

【0023】なお、以上の(1)～(8)の発明における映像制作装置は、(1)～(8)の発明における映像制作方法を実施することを可能としている。

【0024】(9)の発明は、コンピュータに、(1)～(8)のいずれかの発明の映像制作方法を機能させるための処理プログラムを、コンピュータが読み取り可能な媒体に記録した記録媒体であり、(1)～(8)の發

明の映像制作方法ないしそれに準ずるアルゴリズムを記述した処理プログラムやそれに準ずるコンピュータ制御方法を、コンピュータに与え、制御、実行、処理することで、実現することが可能である。なお、上記は、処理プログラムやそれに準ずるコンピュータ制御方法を、コンピュータが実行する際に読み出すことが可能なディスク等やそれに準ずる記録媒体に格納されていることを示す。

【0025】

【発明の実施の形態】以下、図面を用いて本発明の実施の形態を詳しく説明する。

【0026】本発明の第1の実施形態例として、音楽と動画像によって映像を制作する映像制作方法についてのその流れ図を図1に示す。本映像制作方法は、入力される音楽1及び動画像2から、それぞれ音楽特徴点算出工程11及び動画像特徴点算出工程12によってそれぞれの特徴点を算出する。そして、算出された特徴点の一部乃至全部をインデックスとして双方の情報に付加する。これによって、インデックス付き音楽情報3及びインデックス付き動画像情報4を得る。このインデックス付き音楽情報3及びインデックス付き動画像情報4からそれぞれ利用者が、音楽区間指定工程13及び動画像区間指定工程14によって、利用したい音楽区間5と動画像区間6を指定する。それぞれ指定された音楽区間5及び動画像区間6に対して、同期をとって時間長を適合させるため、それぞれのインデックス同士を基準に同期を行う。この際に行われる時間調整工程15では、様々なアプローチがあるが、音楽区間5乃至動画像区間6の一方乃至双方から一部を削除し、時間長をあわせ、最適な切り替えのタイミングで双方の情報区間を同期させることを目的としている。なお、一方的に動画像区間6を削除しても良いし、逆に一方的に音楽区間5を削除しても良い。また、情報区間の時間長に応じて適宜変更しても構わない。時間長を同期をさせるための制御信号として同期信号7を出力後、映像保存表示工程16にて、同期信号7を基に、指定された音楽区間5及び動画像区間6から音楽と動画像からなる映像8を出力する。

【0027】図2は、本実施形態の映像制作方法を実施する映像制作装置の一実施形態例の構成の概要を示すブロック線図であり、21は音楽入力バッファ、22は動画像入力バッファであり、それぞれ音楽1及び動画像2を保管しておくバッファである。なお、音楽1及び動画像2は、あらかじめデジタルデータ化されているものであり、様々な圧縮形式で圧縮されていても良く、ランダムなアクセスが可能であれば良い。したがって本ブロック線図上では、便宜上バッファとしているが、CD-R OMやハードディスク等のデジタルデータを保管する記録メディアであっても良い。

【0028】情報入力バッファに格納された音楽1及び動画像2を、それぞれの特徴的なイベントを表す特徴点

を算出するために、音楽特微量算出部23及び動画像特微量算出部24へ入力する。なお、音楽1を解析して得られる特徴的なイベントとは、音量変化、周波数変化、等であり、動画像2で得られる特徴的なイベントとは、シーン変化点、カメラワーク、動き被写体の有無、動き被写体の領域の大きさや速度等である。音楽特微量算出部23及び動画像特微量算出部24についてはその詳細を後述する。

【0029】音楽特微量算出部23と動画像特微量算出部24にて算出された特徴点をインデクスとして、それぞれの情報に付加することで、インデクス付き音楽情報3及びインデクス付き動画像情報4を得る。音楽と動画像から成る映像の作成にあたっては、利用者が素材として利用したい音楽1と動画像2のうちの一部を選択、指定する必要がある。音楽1と動画像2のうちの一部の指定にあたっては、タイムコード等の時間情報だけを手がかりとして指定しても良いし、撮影した場所や時刻等の情報を手がかりとしても良い。しかし、ある程度の単位にあらかじめ分類されている方が区間単位で選択しやすいため、本実施形態例では、インデクス付き音楽情報3及びインデクス付き動画像情報4のそれぞれのインデクスを分類の手がかりとし、利用者が指定しやすいように、インデクス毎に分類した結果を音楽区間分類表示部25及び動画像区間分類表示部26にて行う。利用者27は、利用者指示部28を通じて、インデクスによって分類表示された動画像及び音楽から映像として利用したい音楽区間5と動画像区間6をそれぞれ音楽区間指定部29及び動画像区間指定部30によって指定する。

【0030】次に、指定された両区間5、6を時間調整部31にて同期を取る。動画像区間5または音楽区間6の一部を除去し、つなげ、同期を行う処理が必要とされる。この詳細な方法については、後述する。時間長調整部31では、音楽区間5及び動画像区間6をどのようなタイミングでそれぞれ再生させると効果的な映像になるを考慮して同期を行う。この再生を行うべきタイミングを時間調整部31から同期信号7として出力する。この同期信号7に基いて、音楽区間5と動画像区間6を互いにどのタイミングで区間の先頭乃至途中から区間の途中乃至最後までを再生させ、メモリ32へ出力する。この出力メモリ32には、テーマ音楽付き映像8が格納されており、利用者27が利用者指示部28を介して行った指示に従って、再生表示部33で映像8を再生したり、ファイル等として保存したりするといった処理を行なう。

【0031】なお、上述の実施形態例で映像素材の指定を行う音楽区間指定部29と動画像区間指定部30において、音楽に時間連續性を持たせた映像を制作したい場合には、指定される音楽区間5は時間的に連續した音楽の一部乃至全部であることが望ましく、動画像区間6は、動画像2から切り出した時間的に異なる短い区間を

複数個指定する方が映像の制作が容易となる。一方で、動画像に連續性を持たせた映像を制作したい場合には、上述の例と全く逆の特性を持つよう区間を指定する方が制作が容易となる。そこで、利用者指示部28で、音楽、動画像のいずれに連續性を持たせた映像を制作したいかという指示を利用者に指定させ、その指定にかなうような区間指定を行わせるような機能を含んでも良い。

【0032】以下、上述した音楽特微量算出部23の詳細について述べ、図3に音楽特微量算出部23を実現する場合のブロック線図の例を示す。この例で、音楽特微量算出部23は、音量極点算出部41と、同一音量継続区間算出部42と、周波数分布算出部43と、周波数分布相関算出部44を具備する。

【0033】音楽特徴点はインデクスとしても用いられ、時間長を調節し同期をとる際の基準点として用いられる。本実施形態例では、音量のパワー変化の際の極大点、極小点、最大点及び最小点45、と、音量が一定している区間である同一音量区間の開始点及び終了点46、周波数分布変曲点47、周波数分布類似区間の開始点、終了点48、を特徴的なイベントが発生した特徴点として用いる。これら特徴点として用いる例を図4に示す。

【0034】それぞれの特徴点算出の具体的な方法を以下に述べる。図4は、この一連の特徴点算出のプロセスを、音声波形101から周波数分布103、周波数分布の個数変化104、周波数分布の相関107を通じて説明している。

【0035】音量変化の極大点、極小点、最大点、最小点については、音量極点算出部41により、音量分布102での音量の変化を追跡し、その変化から算出することができる。

【0036】また、同一音量区間の開始点と終了点については、同一音量継続区間算出部42により、音量分布102での音量変化を追跡し、単位時間当たりの変化の微分値を計算し、微分値が十分0に近い区間がどの程度続くかを観察することで算出することができる。

【0037】周波数変曲点については、周波数分布算出部43により、FFT等によって得られた周波数分布103をサンプリングしたもの（単位時間当たりの分布106）から、高周波数領域である閾値105を上回るパワーをもつ周波数帯域の個数と、低周波領域である閾値105を上回るパワーをもつ周波数帯域との個数とを比較し、それぞれの周波数分布の個数の差分を算出し、個数の+、-の値が変化した点を周波数変曲点としている。この周波数変曲点は、周波数分布103の変化を表していることになり、高周波数成分主体の音楽から低周波数成分主体の音楽への移行やその逆の移行を示し、曲調の変化を反映しやすい。

【0038】周波数分布類似区間の開始点及び終了点については、周波数分布相関算出部44により、単位時間

当たりの周波数分布の相関 106 の高い区間の開始点と終了点としている。分布が類似している区間は同一メロディが連続して何度も繰り返されていることになる重要な区間だと思われるため、その開始点、終了点を特徴点としている。

【0039】次に、上述した動画像特徴点算出部 24 の詳細について述べ、図 5 に動画像特徴点算出工程 12 を実現するための動画像特徴点算出部 24 のブロック線図の例を示す。本例で、動画像特徴点算出部 12 は、シーン変化算出部 51 と、カメラワーク算出部 52 と、動き被写体算出部 53 とを具備する。

【0040】動画像特徴点はインデックスとしても用いることができ、時間長を調節する際に、同期を取る際の基準点や時間短縮のために行う削除を禁止する動画像区間の範囲の選択にも利用される。本実施形態例では、シーン変化点 54 とカメラワークの開始点及び終了点 55 、動き被写体の出現点及び消滅点 56 、動き被写体の領域の大きさと速度を算出する。なお、これら特徴点の例を図 6 に示す。

【0041】それぞれの特徴点の算出するための具体的な方法について述べる。

【0042】シーン変化の検出は、シーン変化算出部 51 により、連続する動画像の類似度を算出し、大きく類似度が変化したところをシーン変化として算出する(図 6 (a))。

【0043】カメラワークに関しては、カメラワーク算出部 52 により、連続する画像中で対応する注目点同士を追跡する。注目点同士の間の距離と方向が画面全体で均一だった場合には、カメラワークがあったと判定する。こうして、カメラワークの有無を判定し、同一シーン内でのカメラワークの開始点および終了点を決定できる(図 6 (b))。

【0044】動き被写体の有無については、動き被写体算出部 53 により、動き被写体領域をカメラワークの有無の抽出方法と同様の方法で抽出する。ただし、カメラワークの場合と動き被写体の場合で異なる点は、画像間で類似している注目点同士の距離と方向の分布が、カメラワークの場合は、画面全体に出現するのに対し、動き被写体の場合には、これらの分布がある時空間の特定の部分に集中する点である。そこで、注目点同士の距離と方向の分布をみて、類似している分布がある時空間に集中的に出現している場合は、動き被写体が存在すると判定する。動き被写体の有無から、同一シーン内の動き被写体出現点及び消滅点を決定することができる(図 6 (c))。また動き被写体の有無が判定された、動き被写体の領域の速度と大きさを特徴量として算出することができる。

【0045】以上のようにして、音楽及び動画像から特徴点を算出し、インデックスとして各情報に付加することで、インデックス付き音楽情報 3 及びインデックス付き動画

像情報 4 を得る。

【0046】次に、インデックス付き音楽情報 3 及びインデックス付き動画像情報 4 から利用者 27 によって指定された音楽区間 5 及び動画像区間 6 を同期させて映像を制作する時間長調整部 31 に関する詳細な説明を加える。

【0047】時間長調整部 31 では、音楽区間と動画像区間の双方ないし一方から区間の一部を除去する処理を行っている。その場合には、音楽を動画像に同期させるために、音楽の方を一方的に除去する、乃至は、その逆など様々な場合を考えることができる。

【0048】ただし、本実施形態例では、映像と音楽を組み合せて制作させる際の最も簡単な実施形態例として、テーマ音楽に複数の動画像を附加させる例を取り上げて、説明を行う。従って、本実施形態例の場合、音楽区間における音楽は、時系列上での連続性を重んじて、除去されることはないものとする一方で、動画像に関しては、適宜音楽との同期を行うため、その一部を除去することをしている。この場合の実施形態例の説明図を図 7 から図 10 で示す。また、説明を簡単に行うため、音楽の連続性を重視するための条件として、音楽区間 5 は、音楽 1 のある連続する区間が選択されているものとし、動画像区間 6 は、動画像 2 の時間的に不連続な複数の区間が選択され、かつ、音楽区間の時間長よりも、動画像区間の時間長の累計の方が短いとする。

【0049】まず、映像の作成の際に用いる素材として、音楽区間 5 と、動画像区間 6 が与えられる。その与えられ方は、一つの音楽に対し、複数の動画像情報、シーン 1 、シーン 2 、…、シーン N までが与えられる。

【0050】それぞれ選択された情報を基に音楽に同期させて動画像の時間長を調整する。図 7 の例では、音楽情報インデックス同士の間にあたる音楽パートの時間長にあわせて、動画像情報の各シーンを同期させている。動画像情報のシーンの時間長が音楽パートよりも長ければ、動画像情報のシーンの一部を削除し、動画像情報のシーンの時間長の方が短ければ、対象としている動画像情報の次のシーンとあわせて同期をとるというルールを用いている。また、カメラワークと動き被写体がある場合には動画像の削除を行わないという禁則処理も行なっている。この条件下での処理プロセスをより詳細に説明する。

【0051】まず、音楽区間の先頭と音楽情報インデックス 1 の間にあたるパート 1 に、複数の動画像区間の先頭区間にあたるシーン 1 を附加させる。このとき、動画像区間のシーン 1 の時間長が対象としている音楽区間の時間長よりも長いので、シーン 1 の一部のみを付加させる。この場合、シーン 1 には特に動画像インデックスがないため、先頭からパート 1 と同じ時間長分だけ切り出す。

【0052】次に、インデックス 1 とインデックス 2 に挟まれたパート 2 に対して、動画像区間を同期させる試みを

行う。この場合、同期に用いる動画像区間にあたるシーン2の時間長は、パート2の時間長よりも短いため、シーン2だけでなく、次のシーン3を用いてパート2に同期させる。シーン2とシーン3をあわせた時間長とパート2の時間長を比較し、映像区間が十分な長さであれば、シーン2の先頭からシーン3の途中のパート2と同じ時間長さ分を切り出す。

【0053】次に、音楽区間のパート3に動画像区間を付加させる。パート3に対応する動画像区間は、図1の場合、それまでの動画像区間の利用から、シーン4が第一の候補となる。この実施形態例では、パート1とシーン1の場合と同様、シーン4の時間長の方がパート3の時間長よりも長いので、シーン4の一部を切り出して、パート3に同期させる。しかし、シーン4には、カメラワークが存在するというインデックスが含まれている。そこで、カメラワークに関わる部分には、時間短縮のための削除を行わないと言う禁則処理を適用し、先頭からカメラワーク以前の前半部分の一部を切り出したシーン4をパート3に同期させる。これは、カメラワークには重要と思われる情報が含まれており、さらにカメラワーク後の画像の方により重要な情報が含まれているといわれていることから、上述の処理を適用する。

【0054】パート4に対するシーン5の同期処理では、カメラワークではなく、動き被写体を含む場合を示している。シーン4と同様に動き被写体を含む動画像部分をシーン5から切り出して、パート4に同期させていく。ただし、動き被写体の場合には、カメラワークの終了時と異なり、動き被写体を含む領域に加えて、その前後時間長が等しくなるような区間を、パート4の時間長分だけシーン5から切り出して、パート4に同期させていている。これにより、音楽区間と同期のとれた映像を得る。

【0055】以上の処理の結果、図7の例では、音楽区間と動画像区間の同期によって映像（動画像情報71）を得ることができた。

【0056】次に、別の例を図8にて示す。これは、音楽区間のパート11、パート12に動画像区間のシーン11を同期させる方法である。この例では、音楽区間のパート11とパート12をあわせた時間長よりも動画像区間シーン11の方の時間長の方が長い場合である。この場合、シーン11に含まれているカメラワークを途中で切断するような同期の方法は行わず、音楽区間のパート11とパート12を含めて、カメラワークを含めた同期を行うこととする。

【0057】また、別の例を図9にて示す。これは、音楽区間のパート21、パート22、パート23に、動画像区間のシーン21、シーン22を同期させる場合である。このときパート21の時間長は、シーン21の時間長よりも長い為、パート21に同期させる動画像区間としてシーン22までを用いる。しかし、シーン22の大

半には動き被写体を多く含み、シーン21とシーン22中の動き被写体部分を含んだ時間長は、パート21の時間長よりも長く、このままでは同期が不可能であるので、図7の例に基いて、パート21とパート22とパート23にシーン21とシーン22を用いて同期を行うことを提案している。なお、それ以上に動画像区間の時間長が、音楽区間の時間長よりも長い場合には、パート21を含めて同期を行うことにしている。すなわち、図9に示したように、音楽区間において必ず同期が取れるよう、音楽区間のパートないし動画像区間のシーンの部分を逐次延長していき、最終的に動画像区間の一部を削ることで、同期がとれる処理を行う。

【0058】以上の図7から図9のように、シーンの一部を削っていくと、想定したよりもシーンの組み合わせによる動画像区間の時間長が短縮されてしまう場合がある。その場合には、シーンの中で最も多く削ったシーンについて、シーンの構成を見直し、削除を行わず、パートの延長を行うことにする。図10に示すように一度削ったものを基に戻して再度検討をする。図10では、パート31、パート32に、シーン31、シーン32をあわせるため、シーン32の一部を除去している。その場合、動画像区間がなくなってしまい、パート33にあわせる動画像区間が存在しないことになってしまった（74の時間長調整中動画像情報）。そのような場合には、パート32にあわせて、シーン32の一部を削ることなく、パート32とパート33とをあわせて、シーン32を削るようにしている（75の時間長調整後動画像情報）。この補正是、全パート及び全シーンが終了した時点での同期がうまく取れなかった場合に行われる。ただし、後述するように、ある音楽区間及びある動画像区間を必ず同期させるような指示が与えられた場合には、その同期情報より前とその同期情報よりも後にかけてそれぞれ時間長調整処理を行う。

【0059】図7から図10で示した例は、映像制作方法の一つの実施形態例であり、その時間調整のルールには他にも様々なバリエーションが考えられる。

【0060】図7から図10の例では、音楽情報インデックスの音量の最大、最小、極小、極大点、同一音量継続区間の開始点、終了点、周波数変曲点の全てのインデックスの全てのインデックスの間に含まれるものとパートとしてみなした。そうではなく、極大点だけを基準としてパートを分割する等の例も考えられる。

【0061】動画像インデックスでは、シーン変化を一つの単位とし、同期の基準とし、動画像情報から必要な情報を抜き出す際に、カメラワークのある部分や動き被写体のある場合には、一部の除去を行わないことにしたが、実際には、この限りでなく、カメラワークや動き被写体のある間でも、一部を削除して構わない。

【0062】また、インデックスとして、上述したインデックス以外を用いても構わない。また、利用者等の指定に

よって特定のパートに特定のシーンを必ず同期させるとといった指定も例としてあげられる、特定のシーンに関しては、パートの時間長に関わらず、その一部を削除し、切り抜いて同期するのではなく、必ず利用するなどの指定が考えられる。

【0063】こうしたルールのバリエーションを持たせることで、より完成度の高い映像を制作することが可能となる。

【0064】なお、上述の映像制作装置において、時間調整部31は、音楽区間と動画像区間を組み合わせている。その際に、音楽区間に基準点を設定し、かつ、動画像区間のあらかじめ指定された条件を除いて削除することにより、音楽区間と動画像区間の同期を行なっている。利用者27が利用者指示部28を介することで、基準点とするインデックスや、削除をおこわない動画像の特定の区間を指定しても良い。一方で、必ず同期させる音楽区間と動画像区間を設定するなどしてもよい。また、自動的な同期が行われた後、その結果に操作を加えるなどの拡張を行っても良い。本実施形態例では、素材とする音楽区間及び動画像区間はそれぞれオリジナルな情報があるとしているが、複数の素材から任意に選択する等の拡張も構わない。

【0065】本発明では、音楽と動画像から映像を自動制作するために、各区間を指定し、時間長調整部にて音楽と同期をとっている。時間長調整部では、双方の同期を行う際に一方の情報から他方の情報の切り替え基準点を設定し、その部分で切り替えるような設定を与えていた。その処理過程において、情報の一部を削除する場合には、その内容を特徴量から推定し、削除することを防ぐといった禁則処理を設けている。

【0066】この時間長調整部における時間同期と禁則処理を応用した他の実施形態例も考えられる。上述の実施形態例では、あらかじめ素材を全て選択した段階で時間長調整部に全素材を投入しているが、逐次素材が投入される毎に最適な時間同期を自動的に行うという逐次時間同期を時間長調整部において適用も可能である。

【0067】例えば、ノンリニア編集機では、音楽、動画像、の適当な区間をG U I（グラフィカル ユーザインターフェース）を通じて逐次配置することで映像を徐々に構築している。そこで、一つ一つの音楽や動画像が与えられた際、与えられた音楽及び動画像の内容を特徴量から推測し、すでに過去に与えられた音楽や動画像の情報とを自動的に時間同期させる。例えば、動画像が与えられていて、さらに音楽が与えられた際に、音楽のメロディやリズムの変わり目を自動検出する。実際に音楽がある時系列上に配置したい際に、既に時系列上に配置された動画像からあらかじめ検出されているシーンチェンジが時系列上の近傍に存在したら、そこにあわせて自動的に配置するとといった応用例も考えられる。

【0068】この実施形態例を、本発明の第2の実施形

態例として以下に説明する。本実施形態例を構成するブロック線図を図11に示す。

【0069】従来のノンリニア編集機と同様に時間情報を基に編集用いる動画像区間5及び音楽区間6を切り出す。次に、それら区間をタイムコードをベースとした編集フィールドに配置する前にそれら区間から特徴的なイベントを算出し、特徴点を算出し、インデックスとして動画像及び音楽に付加して、インデックス付き音楽情報3乃至インデックス付き動画像情報4を得る。

【0070】指定された動画像乃至音楽区間を映像編集フィールド34の利用者の配置したい時系列に配置する。その際に、既存の編集済映像9が映像編集フィールド34にあるため、編集フィールド34では、既存の編集済み映像9にインデックス付き音楽情報3乃至インデックス付き動画像情報4を追加していく。映像編集フィールド34上の編集済映像9は、あらかじめインデックス付き音楽及び動画像から構成された映像であるため、インデックスが存在している。そこで、追記する動画像乃至音楽のインデックスと、既存の編集済み映像上のインデックスとの比較を行う。追記したい音楽乃至動画像にあるインデックスの近傍に、既存の編集済み映像上のインデックスが存在していれば、それらのインデックスが適合するように時間同期を行う。もしインデックス付き動画像区間に、複数の特徴点が存在し、禁則処理にかかるような特徴を持つ区間（例えば、動き被写体が存在する区間）が存在する場合には、既に編集済みの音楽区間に含まれる特徴点と、そうした禁則処理に関わる区間とが同期しないように、自動同期を行う。

【0071】この実施形態例を実装したインターフェース上での例を図12に示す。この例では、あらかじめ動画像と音楽から編集済の映像9が映像編集フィールド34で編集対象となっている。そこへ、新たに動画像区間が指定され、動画像時系列上に挿入を試みる。動画像区間は、あらかじめ指定された時点で動画像特徴点算出部26によって、インデックス付き動画像情報となっているとする。そこで、動画像へ挿入する際に、既に編集済みの音楽情報に特徴点と、挿入したインデックス付き動画像情報の特徴点とを比較し、それらが時系列上である閾値以内に収まっている場合には、自動的に同期を行うように処理する。この同期処理により35に示す映像が得られる。利用者27は、そのような同期が適切であれば、それを基に映像編集を続ければ良いし、同期が不適切であると思えば任意の時系列上に動画像を挿入すれば良い。

【0072】この第2の実施形態例では、編集済みの映像に、動画像区間を新たに編集素材として加える場合に、逐次、時間同期を行う例を示した。もちろん、動画像区間だけでなく音楽区間も同様に扱って良い。また、今回は特徴点としてどのような特徴点を用いるか限定していないが、シーン変化だけといった特徴点に限定を加えてももちろん構わない。

【0073】なお、本発明は、音楽データ及び動画像データを保存し、それらを自由に読み出し可能なハードディスクやそれに準ずる装置と、データを処理する際に必要なバッファやそれに準ずる装置と、最終的に検出されたカット点を表示、出力するディスプレイなどの装置を備え、それらハードディスク、バッファ及びディスプレイなどをあらかじめ定められた手順に基いて制御する中央演算装置などを備えたコンピュータやそれに準ずる装置を基に、上述した実施形態の処理、乃至は、図1から図12までの一連の図に示した方法ないしアルゴリズムを記述した処理プログラムやそれに準ずる物を、該コンピュータに対して与え、制御、実行させることで実現することが可能である。そこで、該処理プログラムやそれに準ずる物が、コンピュータが実行する際に読み出しを実行できるフロッピーディスク(FD)、光磁気ディスク(MO)やそれに準ずる記憶媒体に格納されていても構わない。

【0074】また、本装置を小型のビデオカメラに搭載し、ビデオカメラで撮影したあとすぐに編集し、音楽付きの映像作品を作り出す装置などへの拡張や、ビデオデッキに搭載し、あちらこちらで録画した映像から自分なりの映像作品を作り出す装置などへの拡張等様々な発展、応用が考えられる。

【0075】

【発明の効果】以上説明したように本発明は、音楽を動画像へ同期させる、または、動画像を音楽へ同期させるようにしたので、より効果的に映像を制作することが可能となる。これによって、音楽と動画像を組み合せる音楽付き映像の制作が容易になる。例えば、既存の映像作品に対し、その雰囲気を伝えるテーマ音楽と、内容の一部を抽出し、それらを組み合せることで、短い映像でかつ重要な意味を含み、次回予告や、宣伝、等に利用できる映像を制作することが容易になる効果がある。

【図面の簡単な説明】

【図1】本発明の第1の実施形態例を示す映像制作方法の流れ図である。

【図2】上記第1の実施形態例の映像制作方法を実施するための映像制作装置の実施形態例を示すブロック線図である。

【図3】上記映像制作装置の実施形態例での音楽特徴点算出部の例を示すブロック線図である。

【図4】上記第1の実施形態例での音楽インデックスの例を示す説明図である。

【図5】上記映像制作装置の実施形態例での動画像特徴点算出部の例を示すブロック線図である。

【図6】(a), (b), (c)は、上記第1の実施形態例での動画像インデックスの例を示す説明図である。

【図7】上記映像制作装置の実施形態例での時間長調整

部の時間長調整例(その1)を示す説明図である。

【図8】上記映像制作装置の実施形態例での時間長調整部の時間長調整例(その2)を示す説明図である。

【図9】上記映像制作装置の実施形態例での時間長調整部の時間長調整例(その3)を示す説明図である。

【図10】上記映像制作装置の実施形態例での時間長調整部の時間長調整例(その4)を示す説明図である。

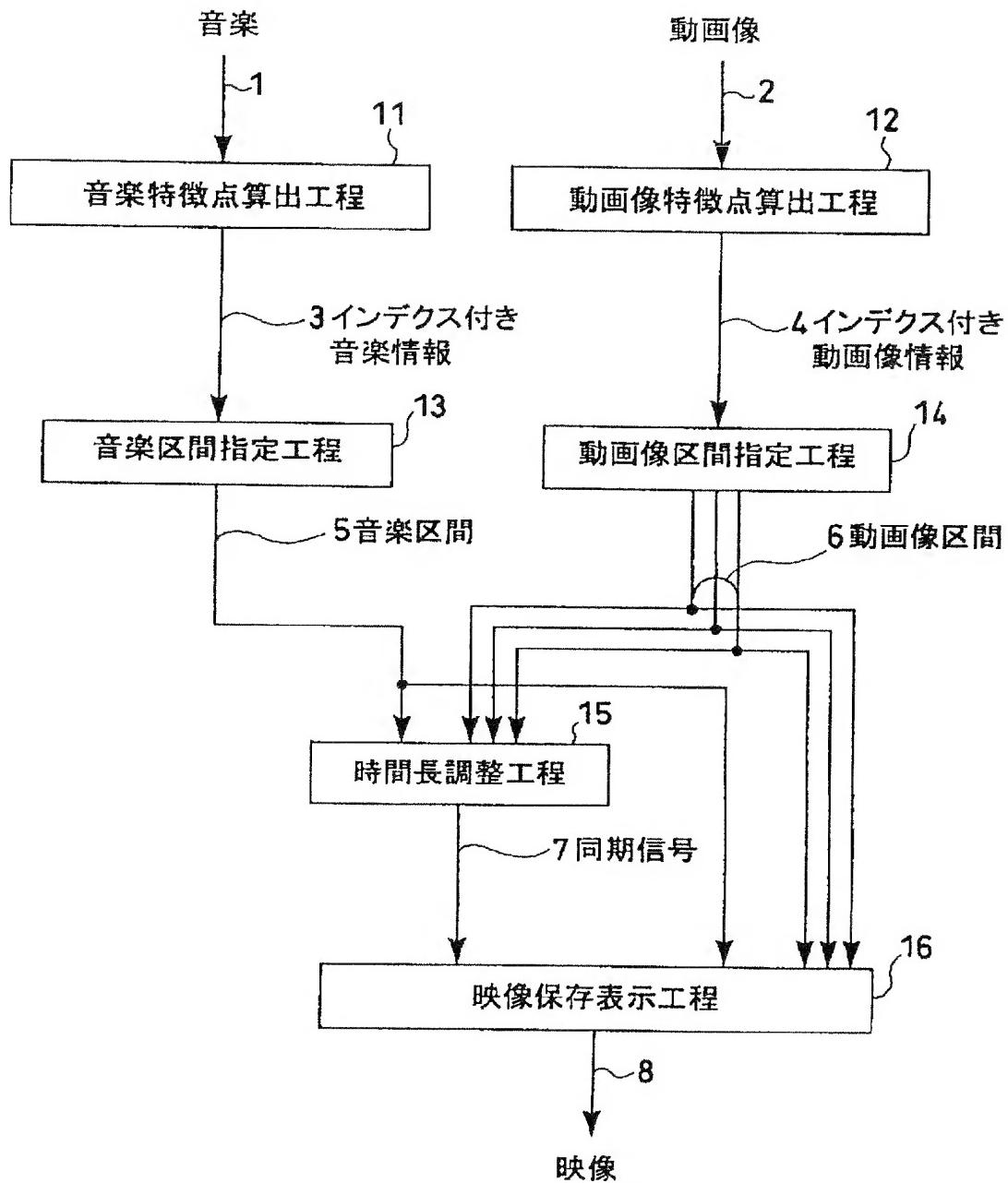
【図11】本発明の第2の実施形態例を示す映像制作装置のブロック線図である。

【図12】上記第2の実施形態例でのインターフェース上の映像編集例を示す模式図である。

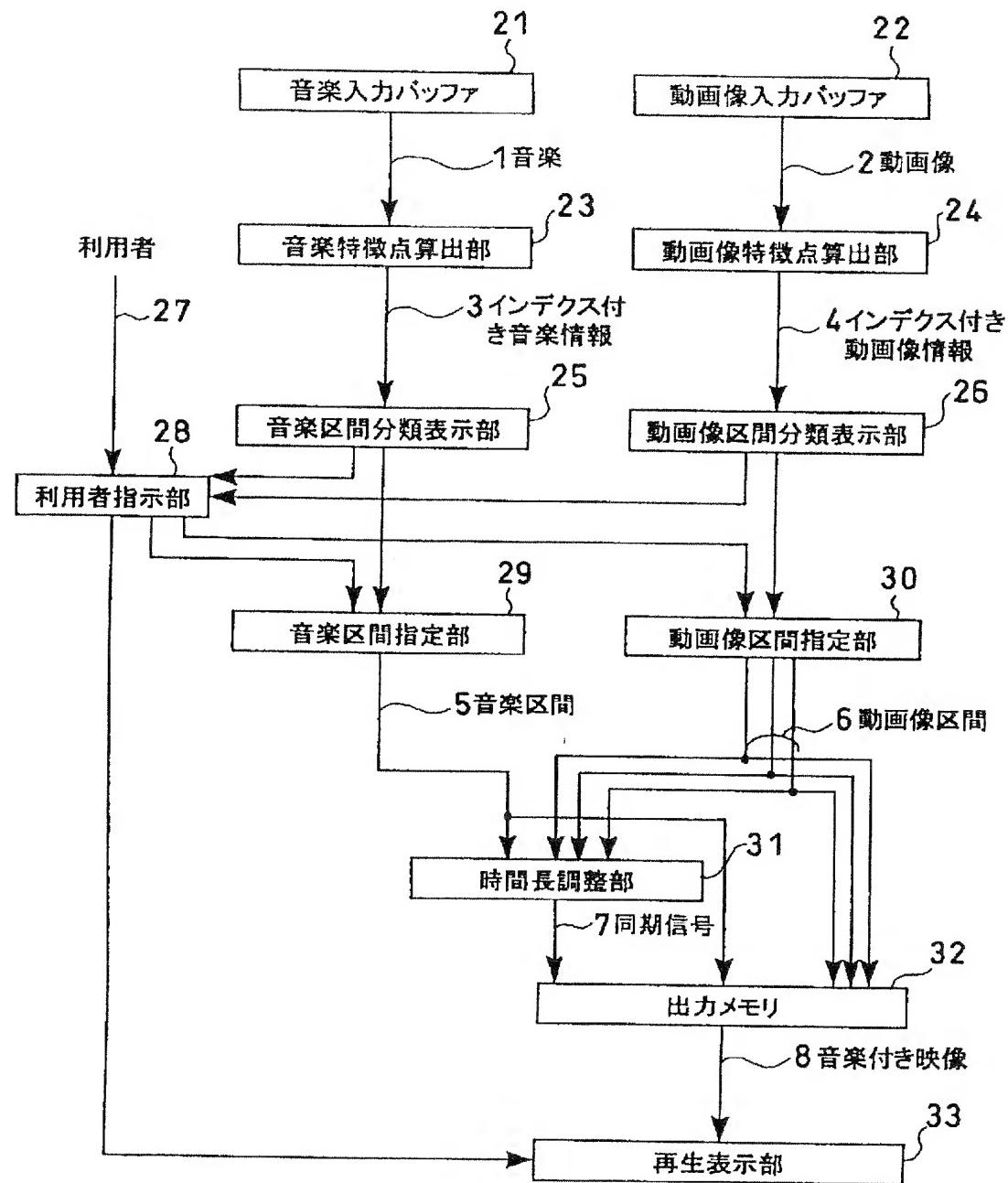
【符号の説明】

- 1…音楽
- 2…動画像
- 3…インデックス付き音楽情報
- 4…インデックス付き動画像情報
- 5…音楽区間
- 6…動画像区間
- 7…同期信号
- 8…音楽付き映像
- 9…編集済映像
- 11…音楽特徴点算出工程
- 12…動画像特徴点算出工程
- 13…音楽区間指定工程
- 14…動画像区間指定工程
- 15…時間調整工程
- 16…映像保存表示工程
- 21…音楽入力バッファ
- 22…動画像入力バッファ
- 23…音楽特徴量算出部
- 24…動画像特徴量算出部
- 25…音楽区間分類表示部
- 26…動画像区間分類表示部
- 27…利用者
- 28…利用者指示部
- 29…音楽区間指定部
- 30…動画像区間指定部
- 31…時間調整部
- 32…出力メモリ
- 33…再生表示部
- 34…映像編集フィールド
- 41…音量極点算出部
- 42…同一音音量継続区間算出部
- 43…周波数分布算出部
- 44…周波数分布相関算出部
- 51…シーン変化算出部
- 52…カメラワーク算出部
- 53…動き被写体算出部

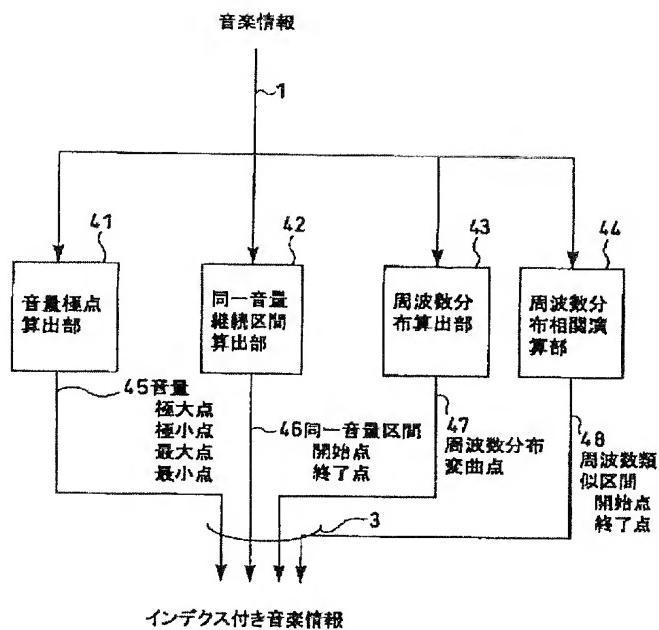
【図1】



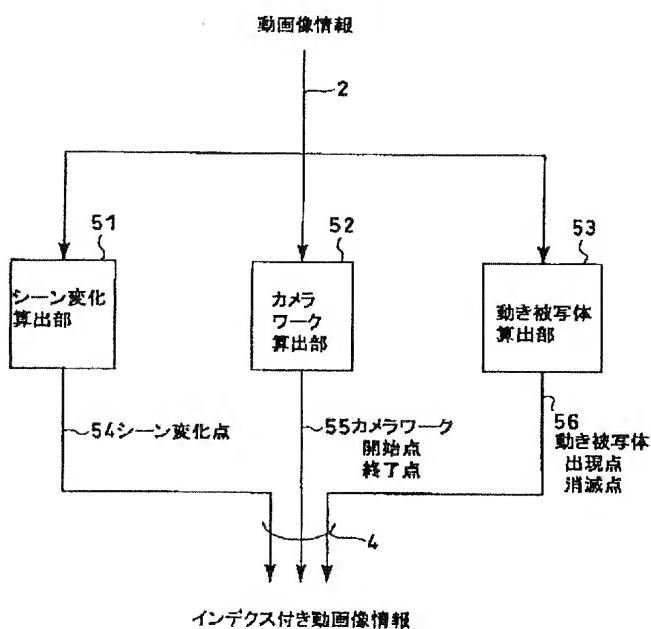
【図2】



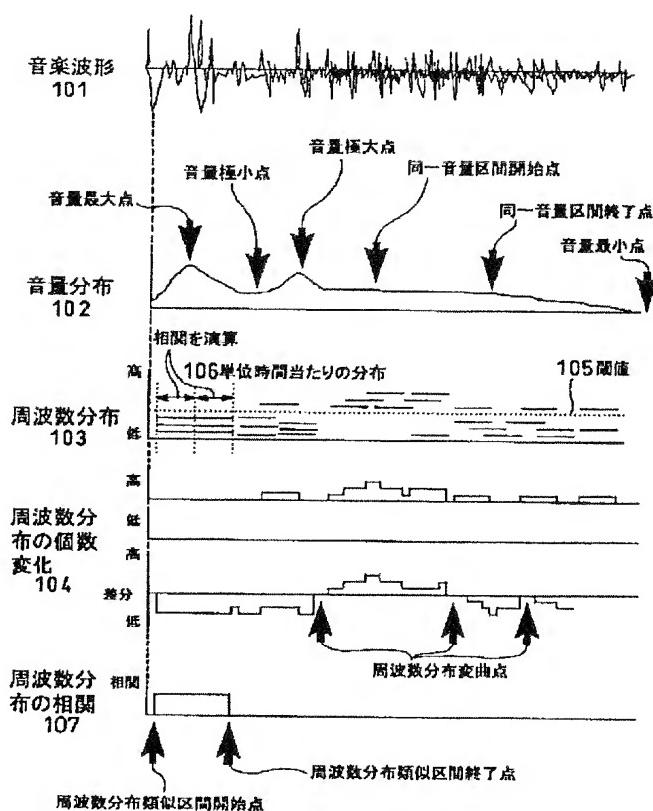
【図3】



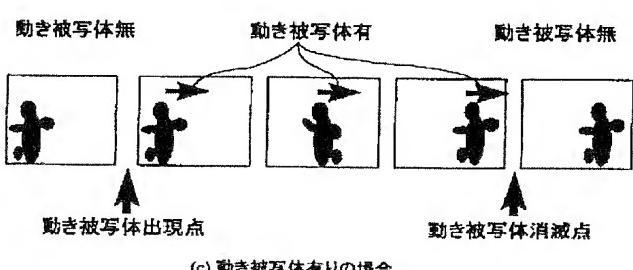
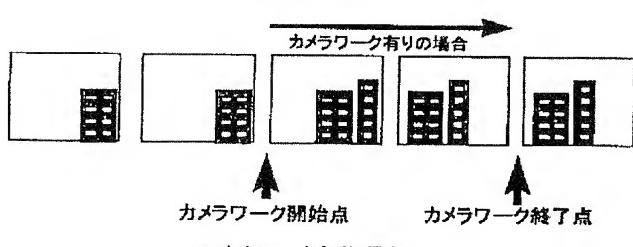
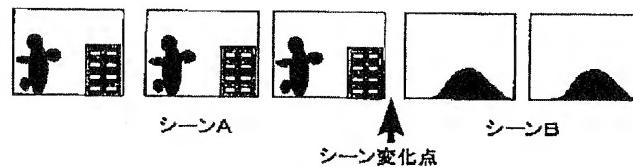
【図5】



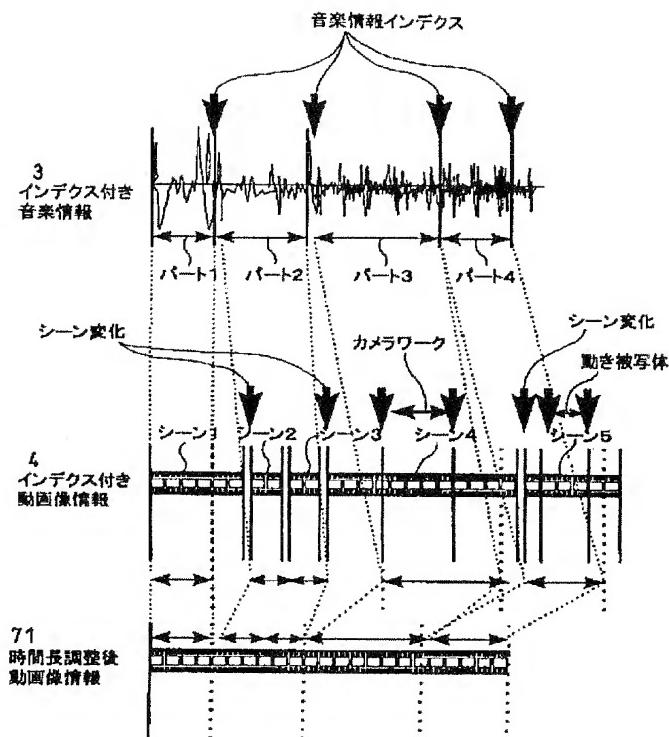
【図4】



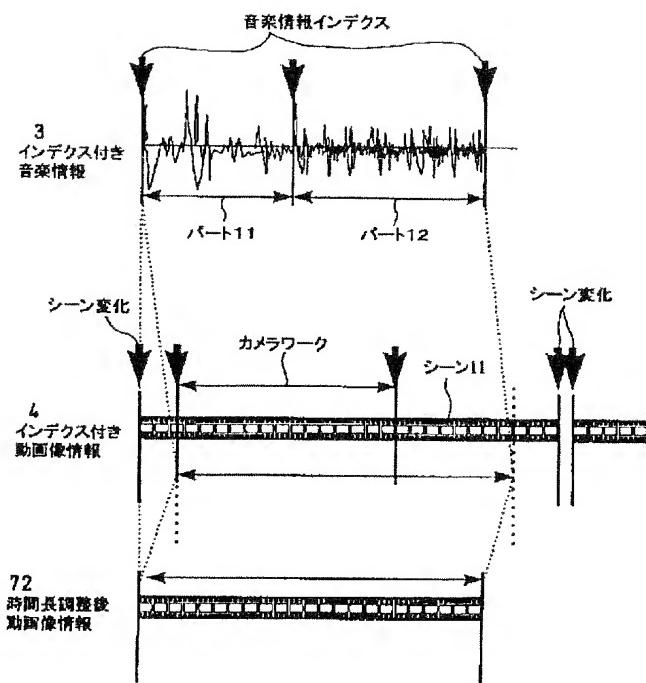
【図6】



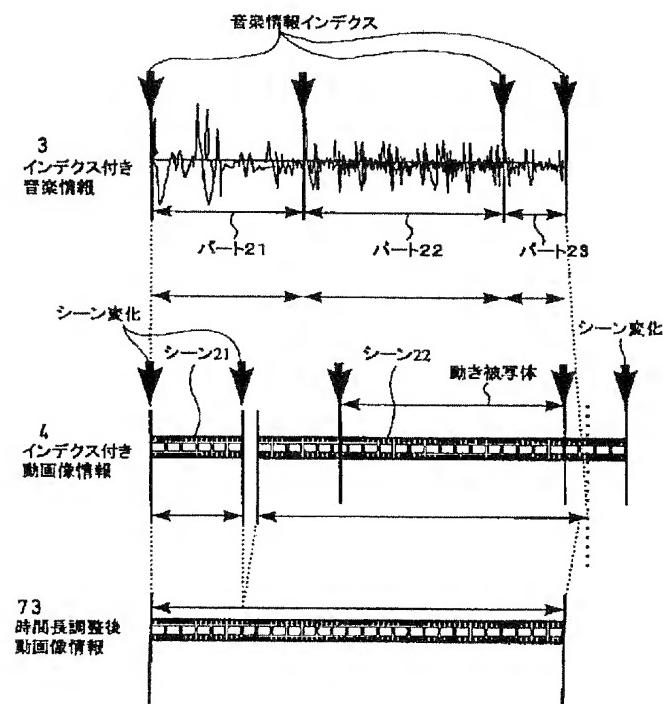
【図7】



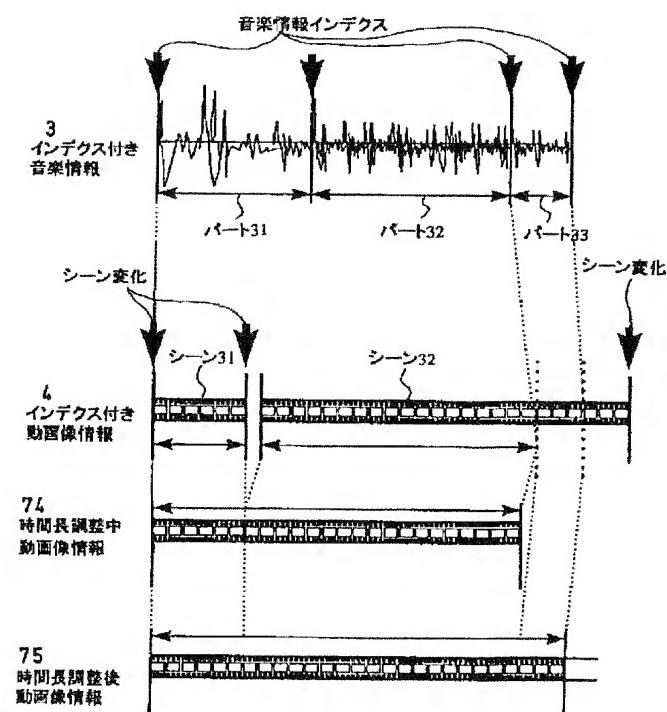
【図8】



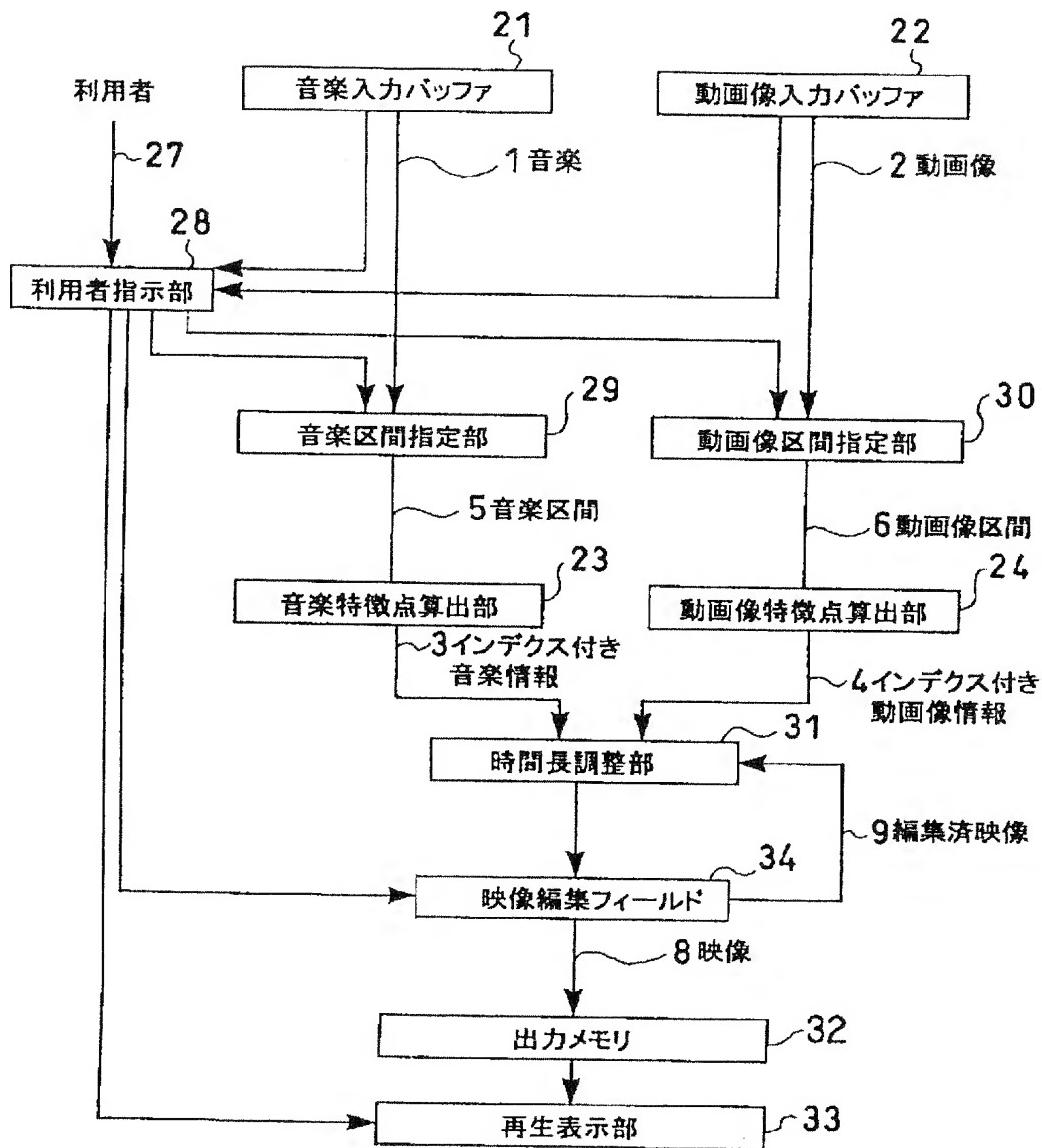
【図9】



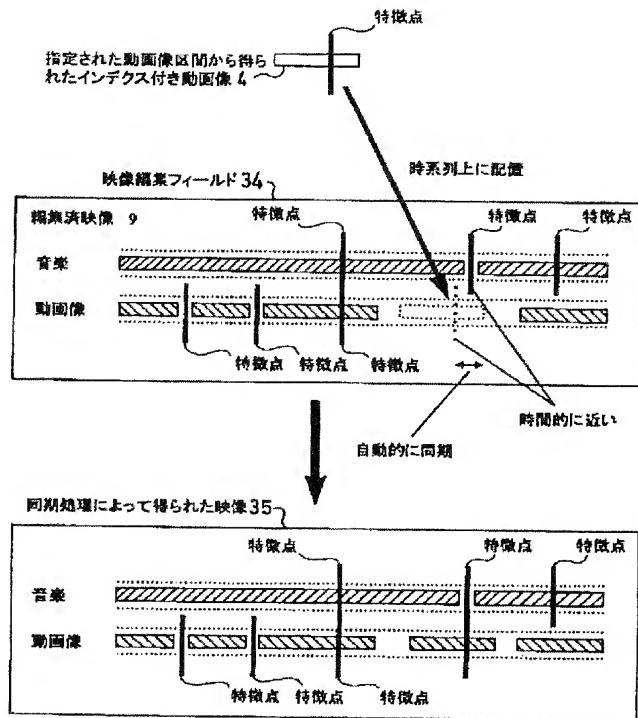
【図10】



【図11】



【図12】



フロントページの続き

(72) 発明者 外村 佳伸
 東京都新宿区西新宿3丁目19番2号 日本
 電信電話株式会社内